The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

INVESTOR IN PEOPLE

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

REC'D 06 OCT 2004
WIPO PCT

GB/04/3391

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed R. Mahoney

Dated 9 September 2004

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
G
28 AUG 2003
NEWPORT

The Patent Office

1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales
NP10 8QQ

28AUG03 E833248-6 D02884
P01/7700 0.00-0320122.5

1.	Your reference	P342351/CPA/MCM
2.	Patent application number *(The Patent Office will fill in this part)*	28 AUG 2003 0320122.5
3.	Full name, address and postcode of the or of each applicant *(underline all surnames)*	Albachem Limited ~~Elvingston Science Centre~~ ~~by Gladsmuir~~ ~~East Lothian~~ ~~EH33 1EH~~ 16 Charlotte Square Edinburgh EH2 4DF
	08641797002 Patents ADP number *(if you know it)*	
	If the applicant is a corporate body, give the country/state of its incorporation	United Kingdom
4.	Title of the invention	"Ligation Method"
5.	Name of your agent *(if you have one)*	Murgitroyd & Company
	"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	Scotland House 165-169 Scotland Street Glasgow G5 8PL
	Patents ADP number *(if you know it)*	1198013 00001198015

		Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*
6.	If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number			

		Number of earlier application	Date of filing *(day / month / year)*
7.	If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application		

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:* — Yes

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	37
Claim *(s)*	7
Abstract	-
Drawing *(s)*	10 fro 16

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*

Request for preliminary examination and search *(Patents Form 9/77)*

Request for substantive examination *(Patents Form 10/77)*

Any other documents *(please specify)*

11. I/We request the grant of a patent on the basis of this application.

Signature Murgitroyd & Co Date 27 August 2003.

Murgitroyd & Company

12. Name and daytime telephone number of person to contact in the United Kingdom — Malcolm C Main — 0141 307 8400

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

Ligation Method

Field of the Invention

This application relates to a method of ligating two or more molecules, for example, small organic molecules, labels, peptides etc. In particular it relates to a method of ligating a peptide, such as ligation of a synthetic peptide to a recombinant peptide.

Background to the Invention

Protein engineering methodologies have proven to be invaluable for generating protein based tools for application in basic research, diagnostics, drug discovery and as protein therapeutics. The ability to manipulate the primary structure of a protein in a controlled manner opens up many new possibilities in the biological and medical sciences. As a consequence, there is a concerted effort on developing methodologies for the site-specific modification of proteins and their subsequent application.

The two main approaches to generating proteins are through recombinant methods or chemical synthesis. To date, the two methods have proved to be complementary; recombinant methodologies enable proteins of any size to be generated but in general they are restricted to the assembly of the proteinogenic amino acids. Thus, in general, the introduction of labels and probes into recombinant proteins has to be implemented post-translationally and does not allow modifications to the protein backbone.

The most common methods for labelling a recombinant protein use an amino or a thiol reactive version of the label that will covalently react with a lysine side chain / N^{α} amino group or a cysteine side chain within the protein respectively. For such labelling methods to be site-specific, an appropriate derivative of the protein must be engineered to contain a unique reactive functionality at the position to be modified. This requires all the other naturally occurring reactive functionalities within the primary sequence to be removed through amino acid mutagenesis. In the case of protein amino functionalities, this is essentially impossible due to the abundance of lysine residues and the presence of the amino functionality at the N-terminus. Likewise, for cysteine this process is laborious and is often detrimental to the function of the protein.

The production of proteins having site-specific modifications and/or labels is more readily achievable using chemical synthesis methods. The chemical synthesis of proteins, however, enables multiple modifications to be incorporated into both side-chain and backbone moieties of the protein in a site-specific manner, but, in general, the maximum size of sequence that can be synthesised and isolated is circa 50 - 100 amino acids.

Protein Ligation

A further approach to the generation of proteins is protein / peptide ligation. In this approach mutually reactive chemical functionalities (orthogonal to the chemistry of the naturally occurring amino acids i.e. which react by mutaually exclusive chemistries compared to the reactions of the reactive moieties of the naturally occuring amino acids) are incorporated at the N- and C-termini of unprotected polypeptide fragments such that when they are mixed, they react in a chemoselective manner to join the two sequences together (Cotton GJ and Muir TW. Chem.Biol., 1999, 6, R247-R254). The principle of chemical ligation is shown schematically in Figure 1.

A number of chemistries have been utilised for the ligation of two synthetic peptides where a diverse range of different chemical functionalities can be incorporated into the termini of polypeptides using solid phase peptide synthesis. These include the reaction between a thioacid and bromo- alkyl to

form a thioester (Schnolzer M and Kent SBH, Science, 1992, 256, 221-225), reaction of an aldehyde with an N-terminal cysteine or threonine to form thiazolidine or oxazolidine respectively (Liu C-F and Tam J P. Proc. Natl. Acad. Sci. USA, 1994, 91, 6584 - 6588), reaction between a hydrazide and an aldehyde to form a hydrazone (Gaertner HF et al, et al Bioconj. Chem., 1992, 3, 262 - 268) reaction of an aminoxy group and an aldehyde to form an oxime (Rose K. *J. Am. Chem. Soc.*, 1994, 116, 30-33), reaction of azides and aryl phosphines to form an amide bond (Staudinger ligation)(Nilsson BL, Kiessling LL, and Raines RT. *Org. Lett.*, 2001, 3, 9-12, Kiick et al *Proc. Natl. Acad. Sci.* USA, 2002, 99, 19-24) , and the reaction of a peptide C-terminal thioester and an N-terminal cysteine peptide to form a native amide bond (Dawson et al. *Science*, 1994, 266, 776) (Native chemical ligation US6184344, EP 0832 096 B1). This method is an extension of studies by Wieland and coworkers who showed that the reaction of ValSPh and CysOH in aqueous buffer yielded the dipeptide ValCysOH (Wieland T et al,. Liebigs Ann. Chem., 1953, 583, 129-149).

Although the native chemical ligation method has proved popular, it requires an N-terminal cysteine and thus, if a cysteine is not present at the appropriate position in the protein, a cysteine needs to be introduced at the ligation site. However, the introduction of extra thiol groups into a protein sequence maybe detrimental to its

structure / function, especially since cysteine has a propensity to form disulfide bonds which may disrupt the folding pathway or compromise the function of the folded protein.

As a consequence of the difficulties and problems associated with known ligation techniques, the ligation of two synthetic fragments generally only enables proteins of circa 100 - 150 amino acids to be chemically synthesised. Although larger proteins have been synthesised by ligating together more than two fragments, this has proved to be technically difficult (Camarero et al. *J. Pept. Res.*, 1998, 54, 303-316, Canne LE et al, *J. Am. Chem. Soc.*, 1999, 121, 8720-8727).

Protein semi-synthesis

Protein ligation technologies that enable both synthetic and recombinantly derived protein fragments to be joined together have been described. This enables large proteins to be constructed from combinations of synthetic and recombinant fragments allowing proteins to be site-specifically modified with both natural and unnatural entities. By utilising such so-called protein semi-synthesis, many different synthetic moieties can be site-specifically incorporated at multiple different sites within a target protein.

In order to utilise recombinant proteins in ligation strategies the recombinant fragments must contain

the appropriate reactive functionalities to facilitate ligation. One approach to introduce a unique reactive functionality into a recombinant protein has been through the periodate oxidation of N-terminal serine containing sequences. Such treatment converts the N-terminal serine into a glyoxyl moiety, which contains an N-terminal aldehyde. Synthetic hydrazide containing peptides have then been ligated to the N-terminus of these protein in a chemoselective manner through hydrazone bond formation with the protein N-terminal aldehyde group (Gaertner HF et al, et al Bioconj. Chem., 1992, 3, 262 - 268, Gaertner HF, et al. *J. Biol. Chem.*, 1994, 269, 7224-7230). Another approach has been to generate recombinant proteins with N-terminal cysteine residues. Synthetic peptides containing C-terminal thioesters have then been site-specifically attached to the N-terminus of these proteins via amide bond formation in a manner analogous to 'native chemical ligation' (Cotton GJ and Muir TW. *Chem. Biol.*, 2000, 7, 253-261). However as with the ligation of synthetic peptides using native chemical ligation techniques, the technology requires a cysteine to be introduced at the ligation site if the primary sequence does not contain one at the appropriate position.

Protein Splicing Techniques

Recently technologies have been developed which enable recombinant proteins containing C-terminal thioester groups to be generated. The C-terminal

thioester functionality provides a unique reactive chemical group within the protein that can be utilised for protein ligation. Recombinant C-terminal thioester proteins are produced by manipulating a naturally occurring biological phenomenon known as protein splicing (Paulus H. *Annu Rev Biochem* 2000, 69, 447-496). Protein splicing is a post-translational process in which a precursor protein undergoes a series of intramolecular rearrangements which result in precise removal of an internal region, referred to as an intein, and ligation of the two flanking sequences, termed exteins (Figure 2). While there are generally no sequence requirements in either of the exteins, inteins are characterised by several conserved sequence motifs and well over a hundred members of this protein domain family have now been identified.

The first step in protein splicing involves an N→S (or N→O) acyl shift in which the N-extein unit is transferred to the sidechain SH or OH group of a conserved Cys/Ser/Thr residue, always located at the immediate N-terminus of the intein. Insights into this mechanism have led to the design of a number of mutant inteins which can only promote the first step of protein splicing (Chong *et al* . *Gene*. 1997, 192, 271-281, (Noren et al., *Angew. Chem. Int. Ed. Engl.*, 2000, 39, 450-466). Proteins expressed as in frame N-terminal fusions to one of these engineered inteins can be cleaved by thiols via an intermolecular transthioesterification reaction, to generate the recombinant protein C-terminal

thioester derivative (Figure 3) (Chong *et al* *Gene.* 1997, 192, 271-281, (Noren et al., *Angew. Chem. Int. Ed. Engl.*, 2000, 39, 450-466)(New England Biolabs Impact System WO 00/18881, WO 0047751). Peptide sequences containing an N-terminal cysteine residue can then be specifically ligated to the C-termini of such recombinant C-terminal thioester proteins (Muir et al *Proc. Natl. Acad. Sci. USA.*, 1998, 95, 6705-6710, Evans Jr et al. *Prot. Sci.*, 1998, 7, 2256-2264) , in a procedure termed expressed protein ligation (EPL) or intein-mediated protein ligation (IPL). As with the previously described ligation techniques, such an approach requires a cysteine to be introduced at the ligation site if one is not suitably positioned with the primary protein sequence and thus is subject to the limitations and associated with the problems of these approaches, such as the potential problems associated with the introduction of an extra thiol group into the primary sequence.

The chemoselective ligation of N-terminal cysteine containing peptides to C-terminal thioester containing peptides, be they synthetic or recombinant, is performed typically at slightly basic pH and in the presence of a thiol cofactor. The strategy also requires a cysteine to be introduced at the ligation site, if one is not suitably positioned within the primary sequence. These requirements of this ligation approach have the potential to alter the structure or function of

both the protein ligation product and the initial reactants.

Protein labelling

Historically protein ligation means the joining together of two peptide / protein fragments but this is synonymous with protein labelling whereby the label is a peptide or derivatised peptide. Equally if a small non-peptidic synthetic molecule contains the necessary reactive chemical functionality for protein ligation, then ligation of the synthetic molecule directly to either the N- or C- termini of the protein affords site-specific labelling of the protein. Thus technologies developed for the ligation of protein fragments can also be used for the direct labelling of either the N- or C- termini of peptides or proteins in a site - specific manner irrespective of their sequence.

Recombinant proteins containing N-terminal glyoxyl functions (generated through periodate oxidation of the corresponding N-terminal serine protein) have been site-specific N-terminally labelled through reaction with hydrazide or aminoxy derivatives of the label (Geoghegan KF and Stroh JG. *Bioconj Chem.*, 1992, 3, 138-146, Alouni S et al. *Eur. J. Biochem.*, 1995, 227, 328 - 334). Also recombinant proteins containing N-terminal cysteine residues have been N-terminally labelled through reaction with thioester containing labels, the label being the acyl substituent of the thioester (Schuler B and Pannell

LK. *Bioconjug. Chem.*, 2002, 13, 1039-43) and aldehyde (Zhao et al. *Bioconj. Chem.*, 1999, 10, 424-430) functionalities to form amides and thiazolidines respectively.

Though a number of methods for ligation of proteins exist each one has its potential drawbacks. There is thus a need for novel ligation methodologies, especially those that are compatible with both synthetic and recombinant fragments, which will complement the existing technologies and add another string to the protein engineers' bow.

Summary of the Invention

The present inventors have overcome a number of problems associated with the prior art and have developed a new method for ligating peptide molecules which overcomes a number of the problems of the prior art.

Accordingly, in a first aspect of the present invention, there is provided a method of producing an oligopeptide product, the method comprising the steps:

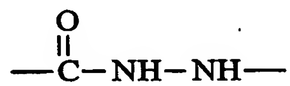
a) providing a first oligopeptide, the first oligopeptide having a reactive moiety,

b) providing a second oligopeptide, the second oligopeptide having a activated ester moiety

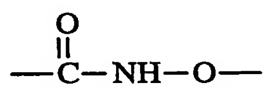
c) allowing the reactive moiety of the first oligopeptide to react with the activated ester moiety of the second oligopeptide to form an

oligopeptide product, in which the first and second oligopeptides are linked via a linking moiety having Formula I, Formula II or Formula III.

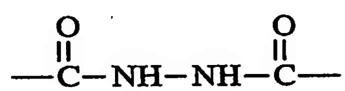
Formula I



Formula II



Formula III



In preferred embodiments, in step (c), where said oligopeptides are linked via a linking moiety having Formula II and where said activated ester moiety of step (b) is not a thioester, said activated ester is a terminal activated ester moiety.

In further preferred embodiments of the invention, said linking moieties are linked via a linking moiety having Formula I or Formula III.

Unless the context demands otherwise, the terms peptide, oligopeptide, polypeptide and protein are used interchangeably.

The activated ester moiety of the first oligopeptide may be any suitable activated ester moiety, such as a thioester moiety a phenolic ester moiety, an hydroxysuccinimide moiety, or an O-acylisourea moiety.

In preferred embodiments of the invention, the activated ester moiety is a thioester moiety. Any suitable thioester peptides may be used in the present invention. In preferred embodiments, the thioester is a thioester wherein the peptide is the acyl substituent of the thioester.

Such thioester peptides may be synthetically or recombinantly produced. The skilled person is well aware of methods known in the art for generating synthetic peptide thioesters. For example, synthetic peptide thioesters may be produced via synthesis on a resin that generates a C-terminal thioester upon HF cleavage (Hojo et al, Bull. Chem. Soc. Jpn., 1993, 66, 2700-2706). Further, the use of 'safety catch' linkers has proved to be popular for generating C-terminal thioesters through thiol induced resin cleavage of the assembled peptide (Shin Y et al, J. Am. Chem. Soc., 1999, 121, 11684-11689).

Moreover, recently technologies have been developed which enable recombinant C-terminal thioester proteins to be generated. Recombinant C-terminal thioester proteins may be produced by manipulating a naturally occurring biological phenomenon known as

protein splicing. As described above, protein splicing is a post-translational process in which a precursor protein undergoes a series of intramolecular rearrangements which result in precise removal of an internal region, referred to as an intein, and ligation of the two flanking sequences, termed exteins.

As described above, a number of mutant inteins which can only promote the first step of protein splicing have been designed (Chong *et al* *Gene*. 1997, 192, 271-281, Noren et al., *Angew. Chem. Int. Ed. Engl.*, 2000, 39, 450-466). Proteins expressed as in frame N-terminal fusions to one of these engineered inteins can be cleaved by thiols via an intermolecular transthioesterification reaction, to generate the recombinant protein C-terminal thioester derivative (Chong *et al* *Gene*. 1997, 192, 271-281, Noren et al., *Angew. Chem. Int. Ed. Engl.*, 2000, 39, 450-466) (New England Biolabs Impact System WO 00/18881, WO 0047751). Such protein thioesters may be used in the methods of the invention (See Figure 3).

Accordingly, in a preferred aspect of the present invention, in step (b), the second oligopeptide is generated by thiol reagent induced cleavage of an intein.

Accordingly, in a second aspect of the present invention, there is provided a method of producing an oligopeptide product, the method comprising the

steps:

a) providing a first oligopeptide, the first oligopeptide having a reactive moiety,

b) (i) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising a second oligopeptide fused N-terminally to an intein domain

(ii) allowing thiol reagent dependent cleavage of the precursor molecule to generate a second oligopeptide molecule, said second oligopeptide molecule having a thioester moiety at its C-terminus

c) allowing the reactive moiety of the first oligopeptide to react with the second oligopeptide molecule to form an oligopeptide product, in which the first and second oligopeptides are linked via a linking moiety having Formula I, II or III.

The reactive moiety of the first oligopeptide may be any suitable reactive moiety. In preferred embodiments of the invention, the reactive moiety is a hydrazine moiety, an amino-oxy moiety or a hydrazide moiety having general formula IV, V or VI respectively.

Formula IV

$$-NH-NH_2$$

Formula V

$$-O-NH_2$$

Formula VI

$$-\overset{\displaystyle O}{\overset{\|}{C}}-NH-NH_2$$

For example, in a particular preferred embodiment, the reactive moiety has Formula IV and, in the oligopeptide product produced by the method of the invention, the first and second oligopeptides are linked via a linking moiety having Formula I.

In a further preferred embodiment, the reactive moiety has Formula V and, in the oligopeptide product produced by the method of the invention, the first and second oligopeptides are linked via a linking moiety having Formula II.

In another preferred embodiment, the reactive moiety has Formula VI and, in the oligopeptide product produced by the method of the invention, the first and second oligopeptides are linked via a linking moiety having Formula III.

As described above, the first oligopeptide comprises a reactive moiety, which, in preferred embodiments, may be a hydrazine moiety (e.g. Formula IV), an amino-oxy moiety (e.g. Formula V) or an hydrazide moiety (e.g. Formula VI).

Hydrazine, hydrazide or aminooxy containing derivatives of synthetic oligopeptides may be readily produced using known methods, for example, solid phase synthesis techniques.

Further, the present inventors have also found that proteins fused N-terminal to an intein domain can be cleaved from the intein by hydrazine treatment in a selective manner to liberate the desired protein as its corresponding hydrazide derivative (for example, see Figure 5).

Accordingly, in further preferred embodiments of the invention, the first oligopeptide is generated by reaction of hydrazine with an oligopeptide molecule comprising the first oligopeptide fused N-terminal to an intein domain.

Indeed the discovery that such protein hydrazides may be produced by means of such a reaction forms an independent aspect of the present invention.

Accordingly, a third aspect of the invention provides a method of generating a protein hydrazide, said method comprising the steps:
(a) providing an protein molecule comprising an oligopeptide fused N-terminal to an intein domain,
(b) reacting said protein molecule with hydrazine, such that the intein domain is cleaved from the oligopeptide to generate a protein hydrazide.

Moreover, as well as using such a reaction to generate a first oligopeptide having a hydrazide moiety at its C-terminal, the first oligopeptide thus being available for reaction with the second oligopeptide having the activated ester moiety, the

present invention further extends to a "one-step" process for ligating two peptides to generate an oligopeptide product.

This may be achieved by reacting a suitable protein linked N-terminal to an intein directly with a polypeptide having a hydrazine, hydrazide or amino-oxy moiety.

Accordingly, in a fourth aspect, the invention provides a method of producing an oligopeptide product, the method comprising the steps:

a) providing a first oligopeptide, the first oligopeptide having a reactive moiety, wherein the reactive moiety is a hydrazine moiety, a hydrazide moiety or an amino-oxy moiety;

(i) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising a second oligopeptide fused N-terminally to an intein domain;

(c) allowing the reactive moiety of the first oligopeptide to react with the precursor oligopeptide molecule to form an oligopeptide product, in which the first and second oligopeptides are linked via a linking moiety having Formula I, Formula II or Formula III.

The ligation technology of the present invention can thus utilise both synthetic and recombinant proteins and peptides. It thus enables the ligation of two or more synthetic, two or more recombinant or a mixture of one or more synthetic with one or more recombinant peptides.

Moreover, as well as providing a novel method of ligating peptides, the present invention may be used for the labelling of synthetic or recombinant peptides.

Accordingly, in a fifth aspect of the present invention, there is provided a method of labelling an oligopeptide, the method comprising the steps:
a) providing a label molecule, the label molecule having a reactive moiety,
b) providing the oligopeptide, the oligopeptide having an activated ester moiety
c) allowing the reactive moiety of the label molecule to react with the activated ester moiety of the oligopeptide to form the labelled oligopeptide, in which the label molecule and the oligopeptide are linked via a linking moiety having Formula I, Formula II or Formula III as defined above,

In preferred embodiments, in step (c), where said label molecule and the oligopeptide are linked via a linking moiety having Formula II and where said activated ester moiety of step (b) is not a thioester, said activated ester is a terminal activated ester moiety.

Alternatively, a label molecule having a terminal activated ester moiety may be used to label an oligopeptide having a reactive moiety. Thus, in a sixth aspect of the invention, there is provided a method of labelling an oligopeptide, the method

comprising the steps:
a) providing a label molecule, the label molecule having an activated ester moiety of which the label is the acyl substituent,
b) providing the oligopeptide, the oligopeptide having a reactive moiety
c) allowing the activated ester moiety of the label molecule to react with the reactive moiety of the oligopeptide to form the labelled oligopeptide, in which the label molecule and the oligopeptide are linked via a linking moiety having Formula I, Formula II or Formula III

wherein, in step (c), where said label molecule and the oligopeptide are linked via a linking moiety having Formula II and where said activated ester moiety of step (b) is not a thioester, said activated ester is a terminal activated ester moiety.

As with the ligation technology, an oligopeptide present as a precursor molecule linked to an intein molecule may be labelled directly. Thus, a seventh aspect of the present invention provides a method of labelling an oligopeptide, the method comprising the steps:
a) providing a label molecule, the label molecule having a reactive moiety,
b) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising an oligopeptide fused N-terminally to an intein domain,
c) allowing the reactive moiety of the label molecule to react with the precursor oligopeptide

molecule to form a labelled oligopeptide product, in which the label molecule and the oligopeptide are linked via a linking moiety having Formula I, Formula II or Formula III as defined above.

The methods of the invention are particularly useful in the ligation of peptides, in particular the ligation of peptides, which, using conventional ligation techniques, would require various protecting groups. The inventors have shown that the methods of the invention may be performed under pH conditions in which only the reactive moieties will react.

In preferred embodiments of the first to seventh aspects of the invention, the method is performed at a pH in the range pH 4.0 to pH 8.5, preferably pH 4.0 to 7.5, more preferably in the range pH 4.5 to pH 7.0, most preferably in the range pH 5.5 to pH 6.5.

For example, the inventors have demonstrated that synthetic peptide C-terminal thioesters specifically react with hydrazine under aqueous conditions at pH 6.0 to form the corresponding peptide hydrazide. This allows ligation methods as described herein to be performed at pH 6.0, without the need for a potentially harmful thiol cofactor (useful if either fragment or final construct is thiol sensitive) and does not lead to the introduction of potentially reactive side-chain groups (such as a thiol) into the protein. Similarly, the inventors have

demonstrated that synthetic peptide C-terminal thioesters specifically react with hydroxylamine under aqueous conditions at pH 6.0 and pH 6.8 to form the corresponding peptide hydroxamic acid.

In an analogous fashion, peptides that contain hydrazine, hydrazide or aminooxy groups can be reacted with thioester derivatives of a label or a peptide to afford site-specific labelling and chemoselective ligation respectively (see, for example, figures 4 and 5).

Furthermore, having demonstrated that recombinant protein hydrazides can be generated by cleavage of protein-intein fusions with hydrazine, the inventors have shown that such protein hydrazides may be ligated by reaction of the hydrazide moiety with reactive groups other than activated ester moieties, for example an aldehyde functionality, a ketone functionality or an isocyanate functionality. This aspect of the invention provides a further novel method of ligating a recombinant peptide to a second peptide or indeed a label.

Thus, an eighth aspect of the invention provides a method of producing an oligopeptide product, the method comprising the steps:

a) providing a first oligopeptide, the the first oligopeptide having an aldehyde or ketone moiety,

b) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising a second oligopeptide fused N-terminally to an intein

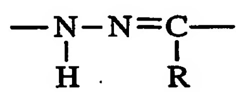
domain,
c) reacting said precursor oligopeptide molecule with hydrazine to generate an oligopeptide molecule comprising an intermediate oligopeptide , said intermediate oligopeptide having a C-terminal hydrazide moiety,
d) allowing the aldehyde or ketone moiety of the first oligopeptide to react with the hydrazide moiety of the intermediate oligopeptide molecule to form an oligopeptide product, in which first oligopeptide and the second oligopeptide are linked via a hydrazone linking moiety.

An example of this aspect is shown in Figure 6.

A ninth aspect of the invention provides a method of labelling an oligopeptide, the method comprising the steps:
a) providing a label molecule, the label molecule having a aldehyde or ketone moiety,
b) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising a first oligopeptide fused N-terminally to an intein domain,
c) reacting said precursor oligopeptide molecule with hydrazine to generate an oligopeptide molecule comprising an intermediate oligopeptide , said intermediate oligopeptide having a terminal hydrazide moiety,
d) allowing the aldehyde or ketone moiety of the label molecule to react with the hydrazide moiety of the intermediate oligopeptide molecule to form a

labelled oligopeptide product, in which the label molecule and oligopeptide are linked via a hydrazone linking moiety.

In preferred embodiments of the eighth and ninth aspects of the invention, the hydrazone moiety has Formula VII:



where R is H or any substituted or unsubstituted, preferably unsubstituted, alkyl group.

In preferred aspects of the eighth and ninth aspects of the invention, the method is performed at a pH in the range pH 1.0 to pH 7.0, preferably pH 1.0 to pH 6.0, more preferably in the range pH 2.0 to pH 5.5, most preferably in the range pH 2.0 to pH 4.5.

In a tenth aspect of the present invention, there is provided an oligopeptide product produced using a method of the invention.

In an eleventh aspect, there is provided a labelled oligopeptide comprising an oligopeptide labelled according to a method of the invention.

Preferred features of each aspect of the invention are as for each of the other aspects mutatis mutandis.

The invention will now be described further in the following non-limiting examples with reference made to the accompanying drawings in which:

Figure 1 illustrates schematically the general principle of chemical ligation.

Figure 2 illustrates schematically the mechanism of protein splicing.

Figure 3 illustrates the generation of recombinant C-terminal thioester proteins.

Figure 4 illustrates ligation of protein and peptide thioesters with hydrazine and aminooxy containing entities, such as labels, peptides and proteins.

Figure 5 illustrates the generation of synthetic and recombinant peptide hydrazides for ligation with thioester containing molecules. Note the peptide or label is is the acyl substituent of the thioester.

Figure 6 illustrates the generation of recombinant peptide hydrazides for ligation with aldehyde and ketone containing molecules.

Figure 7 illustrates SDS-PAGE analysis of Grb2-SH2 - GyrA - CBD (immobilised on chitin beads) treated with DTT and MESNA. Molecular weight markers (lane 1); purified Grb2-SH2 - GyrA - CBD immobilised on chitin beads (lane 4). Grb2-SH2 - GyrA - CBD treated with 100 mM DTT (lanes 5 and 7) or 120 mM MESNA

(lanes 8 and 10). Both the whole reaction slurries (lanes 5 and 8) and the reaction supernatants (lanes 7 and 10) were analysed.

Figure 8 illustrates SDS-PAGE analysis of Grb2-SH2 - GyrA - CBD (immobilised on chitin beads) treated with hydrazine. Molecular weight markers (lane 1); Purified Grb2-SH2 - GyrA - CBD immobilised on chitin beads after 20h treatment with phosphate buffer only (lane 2). Grb2-SH2 - GyrA - CBD treated with 200 mM hydrazine in phosphate buffer for 20 h. The whole reaction slurries were analysed.

Figure 9 illustrates an ESMS spectrum of the C-terminal hydrazide derivative of Grb2-SH2.

Figure 10 shows SDS-PAGE analysis of the reaction between synthetic ketone containing peptide CH3COCO-myc with Grb2-SH2 - C-terminal hydrazide and Cytochrome C. Molecular weight markers (lane 1); Grb2-SH2 - C-terminal DTT thioester (lane 2). Reaction between Grb2-SH2 - C-terminal hydrazide and CH3COCO-myc at time points t=0 h (lane 3), t=24 h (lane 4), t= 48h (lane 5) and t= 72 h (lanes 6). Reaction between Cytochrome C and CH_3COCO-myc at time points t=0 h (lane 7), t=24 h (lane 8), t= 48h (lane 9) and t= 72 h (lanes 10)

Examples

Example 1 -Protein ligation / site specific protein labelling using the reaction of peptide / protein thioesters with compounds containing hydrazine / hydrazide or aminoxy functionalities.

Reaction of a peptide C-terminal thioester with 100mM hydrazine at pH 6.0

200 mM sodium phosphate buffer pH 6.0 containing 100mM hydrazine monohydrate (200 μL) was added to a model synthetic peptide α-thioester termed AS626p1A (200 μg) to yield a final peptide concentration of 317 μM. AS626p1A has sequence ARTKQ TARK(Me)$_3$ STGGKAPRKQ LATKAARK-COS-$(CH_2)_2$-$COOC_2H_5$ (SEQ ID NO: 1) wherein a single Alanine residue (which may be any one of the Alanine residues of SEQ ID NO: 1) is substituted by an Arginine residue. The reaction was incubated at room temperature and monitored with time by analytical reversed phase HPLC. Vydac C18 column (5 μM, 0.46 x). Linear gradients of acetonitrile water / 0.1% TFA were used to elute the peptides at a flow rate of 1 mL min^{-1}. Individual peptides eluting from the column were characterised by electrospray mass spectrometry.

Reaction of a peptide C-terminal thioester with 100mM hydroxylamine at pH 6.0

200 mM sodium phosphate buffer pH 6.0 containing 100mM hydroxylamine hydrogen chloride (200 μL) was added to AS626p1A (200 μg) to yield a final peptide concentration of 317 μM. The reaction was incubated at room temperature and monitored with time by

analytical reversed phase HPLC. Vydac C18 column (5 μM, 0.46 x). Linear gradients of acetonitrile water / 0.1% TFA were used to elute the peptides at a flow rate of 1 mL min^{-1}. Individual peptides eluting from the column were characterised by electrospray mass spectrometry.

Reaction of a peptide C-terminal thioester with 100 mM hydroxylamine at pH 6.8

200 mM sodium phosphate buffer pH 6.8 containing 100mM hydroxylamine hydrogen chloride (200 μL) was added to AS626p1A (200 μg) to yield a final peptide concentration of 317 μM. The reaction was incubated at room temperature and monitored with time by analytical reversed phase HPLC. Vydac C18 column (5 μM, 0.46 x). Linear gradients of acetonitrile water / 0.1% TFA were used to elute the peptides at a flow rate of 1 mL min^{-1}. Individual peptides eluting from the column were characterised by electrospray mass spectrometry.

Reaction of a peptide C-terminal thioester with 10mM hydroxylamine at pH 6.8

See above procedure.

Reaction of a peptide C-terminal thioester with 10mM hydroxylamine at pH 7.5

See above procedure.

Reaction of a peptide C-terminal thioester with 2mM hydroxylamine at pH 7.5

See above procedure.

Results

These examples demonstrate the novel strategy for protein ligation / site specific protein labelling of both synthetic and recombinant protein sequences of the invention using the reaction of peptide / protein C-terminal thioesters with compounds containing hydrazine / hydrazide or aminoxy functionalities.

As described above, a purified synthetic 27 amino acid α-thioester peptide (the ethyl 3-mercaptopropionate thioester derivative) was treated with hydrazine and hydroxylamine under various conditions (Table 1).

Treatment with 100 mM hydrazine at pH 6.0 formed a peptide species that eluted earlier than the starting thioester peptide as analysed by HPLC. This material was identified as the expected peptide hydrazide by ESMS: observed mass = 3054 Da, expected (av. isotope comp) 3053 Da. The reaction of the peptide C-terminal thioester with hydrazine to form the peptide hydrazide was monitored with time by reverse phase HPLC. Only the desired material was formed with no side product formation even after 3 days. The stability of the peptide hydrazide, under the reaction conditions, indicates that the reaction occurs at the C-terminal thioester moiety and is chemoselective in nature. It also highlights the applicability of this reaction for protein ligation

and labelling. (2 h 70% conversion , 4h 95% conversion)

To ascertain whether aminooxy containing compounds chemoselectively react with peptide / protein C-terminal thioesters, to afford protein ligation and site-specific labelling, a synthetic C-terminal thioester peptide was treated with hydroxylamine under various conditions (Table 1).

A purified synthetic 27 amino acid C-terminal thioester peptide (ethyl 3-mercaptopropionate thioester, observed mass 3155 Da) was incubated at room temperature with different hydroxylamine concentrations in aqueous buffers of varying pH. In all cases the peptide C-terminal thioester reacted to form a single product that eluted earlier than the starting thioester peptide as analysed by reverse phase HPLC. This material corresponds to the expected hydroxamic acid peptide as determined by ESMS: observed mass = 3052 Da, expected (av. isotope comp) 3054 Da. The kinetics of the reaction were monitored using reverse phase HPLC. The peptide C-terminal thioester is converted to the corresponding peptide hydroxamic acid in a clean fashion with no side-product formation. As expected increasing the pH of the reaction buffer accelerates the rate of reaction. With 100mM NH_2OH on moving from pH 6.0 to pH 6.8 the percentage product formation after 1h increases from 25% to 91%. The rate of reaction with 100 mM NH_2OH pH 6.0 is comparable with 10 mM NH_2OH at pH 6.8.

The rate of reaction of the peptide C-terminal thioester with hydroxalymine, to form the corresponding hydroxamic acid, increases with increasing pH and decreases with decreasing NH_2OH concentrations. To identify conditions of pH and reactant concentration suitable for peptide / protein labelling and ligation, the labelling was performed under increasing pH and decreasing NH_2OH concentrations.

The reaction with 10 mM was 83% complete after 4h at pH 6.8, while at pH 7.5 it was 83% complete after 2h. On further decreasing the NH_2OH concentration to 2 mM the reaction rate at pH 7.5 decreased markedly, 70% of the starting peptide α-thioester being converted to the corresponding hydroxamic acid after 8hrs. It was noted that a small amount of a side-product corresponding in mass to the peptide acid was formed during the reaction. Presumably this is formed by a competing hydrolysis side reaction at pH 7.5, which was not observed with 10 mM NH_2OH at pH 7.5 due to the faster reaction at this higher reactant concentration.

Reactant	Concentration	pH	Percentage product formation with time				
			1hr	2hr	4hr	8hr	72hr
NH_2NH_2	100 mM	6.0	-	70	100		
NH_2OH	100 mM	6.0	25	48.1	76.3	-	100
NH_2OH	100 mM	6.8	91	100			
NH_2OH	10 mM	6.8	26	-	83	100	
NH_2OH	10 mM	7.5	-	82.7	100	100	
NH_2OH	2 mM	7.5	11.2	17	38	70	80*

Table 1

*All starting material has reacted with 80% conversion to the desired product and ~20% to the hydrolysis side-product.

Example 2- Generation of recombinant C-terminal hydrazide proteins through the selective cleavage of protein - intein fusions with hydrazine, and their subsequent use in ligation / labelling reactions.

To investigate (i) the ability to generate recombinant C-terminal hydrazide proteins through the selective cleavage of protein - intein fusions with hydrazine, and (ii) their subsequent use in ligation / labelling reactions, the SH2 domain of the adapter protein Grb2 was chosen as a model system.

Sequence of human Grb2 SH2 domain
HPW FFGKIPRAKA EEMLSKQRHD GAFLIRESES APGDFSLSVK
FGNDVQHFKV LRDGAGKYFL WVVKFNSLNE LVDYHRSTSV
SRNQQIFLRD IEQVPQQPT

Expression of Grb2-SH2 domain - GyrA intein fusion.

The DNA sequence encoding the SH2 domain of human Grb2 appended at its C-terminus with an extra glycine residue was cloned into the pTXB1 expression plasmid (NEB). This vector $pTXB1_{Grb2\text{-}SH2\ (Gly)}$ encodes for a fusion protein whereby the SH2 domain of Grb2 is linked via a glycine residue to the N-terminus of the GyrA intein, which is in turn fused to the N-terminus of a chitin binding domain region (CBD). *E. coli cells* were transformed with this plasmid and grown in LB medium to mid log phase and protein expression induced for 4h at 37°C with 0.5 mM IPTG. After centrifugation the cells were re-suspended in lysis buffer (0.1 mM EDTA, 250 mM NaCl, 5% glycerol, 1 mM PMSF, 25 mM HEPES, pH 7.4) and lysed by sonication. The soluble fraction was loaded onto a chitin column pre- equilibrated in lysis buffer. The column was then washed with wash buffer (1 mM EDTA, 250 mM NaCl, 0.1% Triton-X 100, , 25 mM HEPES, pH 7.0) to yield purified Grb2-SH2 - GyrA-CBD immobilised on chitin beads (Figure 7).

Generation of Grb2-SH2 C-terminal thioesters by thiol induced cleavage of the Grb2-SH2 - GyrA intein fusion.

To ascertain that the intein domain within the protein was functional the fusion protein was exposed to thiols to assess the extent of cleavage via transthioesterification. Chitin beads containing immobilised Grb2-SH2 - GyrA-CBD were equilibrated into 200 mM NaCl, 200 mM phosphate buffer pH 7.4. Dithiothreitol (DTT) or 2-mercaptoethanesulfonic

acid (MESNA) were then added to the beads in 200 mM NaCl, 200 mM phosphate buffer pH 7.4 to give a 50% slurry with a final thiol concentration of 100 mM or 120 mM respectively. The mixtures were then rocked at room temperature and aliquots analysed by SDS-PAGE. After 48 hours the supernatants from the reactions were isolated and subsequently analysed by HPLC and ESMS.

Treatment of Grb2-SH2 – GyrA intein – CBD fusion with both DTT and MESNA resulted in cleavage of the fusion protein into two protein species (Figure 7). The molecular size of the two fragments corresponds to that of the Grb2 – SH2 and the GyrA – intein fusion, indicative that cleavage has taken place at the SH2 – intein junction. Cleavage of the precursor fusion protein liberated the SH2 domain into the supernatant while the GyrA intein-CBD portion remained immobilized on the chitin beads. After cleavage with both DTT or MESNA, ESMS analysis of the supernatants confirmed that the Grb2-SH2 was generated as either the expected DTT or MESNA C-terminal thioester derivatives respectively.

Expected mass of Grb2-SH2 DTT - C-terminal thioester = 12173.9 Da; observed mass 12173.5 Da. Expected mass of Grb2-SH2 MESNA - C-terminal thioester = 12162.0 Da; observed mass 12163.0 Da.

Generation of Grb2-SH2 C-terminal hydrazide by hydrazine induced cleavage of the Grb2-SH2 - GyrA intein fusion.

The thioester linkage between Grb2-SH2 and the GyrA intein in the precursor fusion protein is expected to be cleaved with hydrazine, the chemoselective reaction of hydrazine, at the thioester moiety, liberating Grb2-SH2 domain into the supernatant as its corresponding C-terminal hydrazide derivative. Chitin beads containing immobilised Grb2-SH2 - GyrA-CBD were therefore equilibrated into 200 mM NaCl, 200 mM phosphate buffer pH 7.4 and hydrazine monohydrate added in the same buffer to give a 50% slurry with a final hydrazine concentration of 200 mM. The mixture was then rocked at room temperature and analysed by SDS-PAGE (Figure 8). After 20 hours the supernatant was removed and analysed by HPLC and ESMS.

Treatment of Grb2-SH2 - GyrA intein - CBD fusion with hydrazine resulted in cleavage of the fusion protein into two species. The molecular size of the two fragments as analysed by SDS-PAGE corresponded to Grb2 - SH2 and the GyrA - intein fusion, indicative that cleavage has taken place at the unique thioester linkage between the SH2 - intein domains. Cleavage of the precursor fusion protein liberates the SH2 domain into the supernatant while the GyrA intein-CBD portion remained immobilized on the chitin beads. HPLC and ESMS analysis of the cleavage supernatant confirmed that a single protein species was generated that corresponds to the C-terminal hydrazide derivative of Grb2-SH2. Expected mass of Grb2-SH2 C-terminal hydrazide = 12051.7 Da; observed mass 12053.0 Da. (Figure 9)

After 20 h of reaction Grb2-SH2 C-terminal hydrazide was isolated form the supernatant using RPHPLC and lyophilised.

Ligation of aldehyde and ketone containing peptides and labels to recombinant C-terminal hydrazide containing proteins.

It was anticipated that recombinant protein C-terminal hydrazides, generated by hydrazine treatment of the corresponding intein fusion precursor, can be site-specifically modified by chemoselective ligation with aldehyde and ketone containing peptides and labels. To demonstrate such an approach the ability of a synthetic ketone containing peptide to ligate with the Grb2-SH2 C-terminal hydrazide generated above was investigated. A synthetic peptide corresponding to the c-myc epitope sequence was synthesised GEQKLISEEDL-NH_2 whereby pyruvic acid was coupled to the amino terminus of the peptide as the last step of the assembly. This peptide (designated CH_3COCO-myc) was purified to > 95% purity by RPHPLC and lyophilised (ESMS expected monoisotopic mass 1328.6 Da; observed mass 1328.6 Da).

A sample of CH_3COCO-myc peptide was dissolved in 100 mM sodium acetate buffer pH 4.5 to give a 4 mM peptide concentration. This peptide solution (100 µL) was then added to an aliquot of lyophilised Grb2-SH2 C-terminal hydrazide protein (~ 250 µg) and the reaction monitored by SDS-PAGE (Figure 10) As a

control CH_3COCO-myc was also incubated with Cytochrome C, a protein of similar same size to Grb2-SH2 but absent of a hydrazide functionality.

SDS-PAGE analysis showed that CH_3COCO-myc peptide has indeed ligated with Grb2-SH2 C-terminal hydrazide as indicated by the conversion of Grb2-SH2 C-terminal hydrazide into a protein species of a higher molecular weight (approximately 1000-2000 Da higher). The reaction was virtually complete after 24 h and the reaction product appeared to be stable. On the other hand there was no observable change to Cytochrome C with time i.e no ligation, establishing that the ligation reaction was occurring at the C-terminal hydrazide functionality of Grb2-SH2.

After 96 h of reaction the product from the Grb2-SH2 ligation reaction was isolated by HPLC and characterised by ESMS. Chemoselective ligation of CH_3COCO-myc to Grb2-SH2 C-terminal hydrazide via hydrazone bond formation would give a product of expected mass 13363.7 Da. The observed product mass was 13364.1 Da indicating that the desired ligation product had been formed.

In summary, the present invention provides novel methods of protein ligation that enable both synthetic and recombinantly derived protein fragments to be efficiently joined together in a regioselective manner. This thus enables large proteins to be constructed from combinations of synthetic and recombinant fragments and allows proteins of any size to be site-specifically

modified in an unprecedented manner. This is of major importance for biological and biomedical science and drug discovery when one considers that the ~ 30,000 human genes yield hundreds of thousands of different protein species through post-translational modification. Such post-translationally modified proteins cannot be accessed through current recombinant technologies.

The application of such protein ligation techniques may be used for protein based tools, protein therapeutics and in *de novo* design and may open up many new avenues in biological and biomedical sciences that have hitherto not been possible.

All documents referred to in this specification are herein incorporated by reference. Various modifications and variations to the described embodiments of the inventions will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes of carrying out the invention which are obvious to those skilled in the art are intended to be covered by the present invention.

Claims

1. A method of producing an oligopeptide product, the method comprising the steps:

a) providing a first oligopeptide, the first oligopeptide having a reactive moiety,

b) providing a second oligopeptide, the second oligopeptide having a activated ester moiety

c) allowing the reactive moiety of the first oligopeptide to react with the activated ester moiety of the second oligopeptide to form an oligopeptide product, in which the first and second oligopeptides are linked via a linking moiety having Formula I, Formula II or Formula III.

Formula I

$$-\overset{\overset{\displaystyle O}{\|}}{C}-NH-NH-$$

Formula II

$$-\overset{\overset{\displaystyle O}{\|}}{C}-NH-O-$$

Formula III

$$-\overset{\overset{\displaystyle O}{\|}}{C}-NH-NH-\overset{\overset{\displaystyle O}{\|}}{C}-$$

2. The method according to claim 1 wherein the terminal activated ester moiety is a thioester wherein the peptide is the acyl substituent of

the thioester.

3.The method according to claim 2, wherein said second polypeptide is generated by thiol reagent dependent cleavage of a precursor molecule, said precursor molecule comprising a second oligopeptide fused N-terminally to an intein domain.

4. A method of producing an oligopeptide product, the method comprising the steps:

a) providing a first oligopeptide, the first oligopeptide having a reactive moiety,

(i) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising a second oligopeptide fused N-terminally to an intein domain

(ii) allowing thiol reagent dependent cleavage of the precursor molecule to generate a second oligopeptide molecule, said second oligopeptide molecule having a thioester moiety at its C-terminus,

c) allowing the reactive moiety of the first oligopeptide to react with the second oligopeptide molecule to form an oligopeptide product, in which the first and second oligopeptides are linked via a linking moiety having Formula I, II or III.

5. The method according to any one of the preceding claims wherein the reactive moiety is a hydrazine moiety, a hydrazide moiety or an aminooxy moiety.

6. The method according to claim 5, wherein said first oligopeptide is produced by reaction of

hydrazine with a precursor molecule, said precursor molecule comprising a precursor oligopeptide fused N-terminally to an intein domain via a thioester moiety.

7. A method of producing an oligopeptide product, said method comprising the steps:
a) providing a first oligopeptide, the first oligopeptide having a reactive moiety, wherein the reactive moiety is a hydrazine moiety, a hydrazide moiety or an amino-oxy moiety;
(i) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising a second oligopeptide fused N-terminally to an intein domain;
(c) allowing the reactive moiety of the first oligopeptide to react with the precursor oligopeptide molecule to form an oligopeptide product, in which the first and second oligopeptides are linked via a linking moiety having Formula I, Formula II or Formula III.

8. The method according to any one of the preceding claims, wherein the first oligopeptide or the second oligopeptide is a recombinant oligopeptide and the other of the the first oligopeptide and the second oligopeptide is a synthetic polypeptide.

9. The method according to any one of claims 1 to 7, wherein the first oligopeptide and the second

oligopeptide are recombinant oligopeptides.

10. The method according to any one of claims 1 to 7, wherein the first oligopeptide and the second oligopeptide are synthetic oligopeptides.

11. A method of generating a protein hydrazide, said method comprising the steps:
(a) providing a protein molecule comprising an oligopeptide fused N-terminal to an intein domain,
(b) reacting said protein molecule with hydrazine, such that the intein domain is cleaved from the oligopeptide to generate a protein hydrazide.

12. The method according to any one of the preceding claims wherein the method is performed at a pH in the range pH 5.5 to 7.5.

13. A method of producing an oligopeptide product, the method comprising the steps:
a) providing a first oligopeptide, the the first oligopeptide having an aldehyde or ketone moiety,
b) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising a second oligopeptide fused N-terminally to an intein domain,
c) reacting said precursor oligopeptide molecule with hydrazine to generate an oligopeptide molecule comprising an intermediate oligopeptide , said intermediate oligopeptide having a

terminal hydrazide moiety,
d) allowing the aldehyde or ketone moiety of the first oligopeptide to react with the hydrazide moiety of the intermediate oligopeptide molecule to form an oligopeptide product, in which first oligopeptide and the second oligopeptide are linked via a hydrazone linking moiety.

14. An oligopeptide product produced by the method of any one of the preceding claims.

15. A method of labelling an oligopeptide, the method comprising the steps:
a) providing a label molecule, the label molecule having a reactive moiety,
b) providing the oligopeptide, the oligopeptide having a activated ester moiety
c) allowing the reactive moiety of the label molecule to react with the activated ester moiety of the oligopeptide to form the labelled oligopeptide, in which the label molecule and the oligopeptide are linked via a linking moiety having Formula I, Formula II or Formula III.

16. The method according to claim 15, wherein in step (c), where said label molecule and the oligopeptide are linked via a linking moiety having Formula II and where said activated ester moiety of step (b) is not a thioester, said activated ester is a terminal activated ester moiety.

17. A method of labelling an oligopeptide, the method comprising the steps:
a) providing a label molecule, the label molecule having an activated ester moiety of which the label is the acyl substituent,
b) providing the oligopeptide, the oligopeptide having a reactive moiety
c) allowing the activated ester moiety of the label molecule to react with the reactive moiety of the oligopeptide to form the labelled oligopeptide, in which the label molecule and the oligopeptide are linked via a linking moiety having Formula I, Formula II or Formula III, wherein, in step (c), where said label molecule and the oligopeptide are linked via a linking moiety having Formula II and where said activated ester moiety of step (b) is not a thioester, said activated ester is a terminal activated ester moiety.

18. A method of labelling an oligopeptide, the method comprising the steps:
a) providing a label molecule, the label molecule having a reactive moiety,
b) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising an oligopeptide fused N-terminally to an intein domain,
c) allowing the reactive moiety of the label molecule to react with the precursor oligopeptide molecule to form a labelled oligopeptide product, in which the label molecule and the oligopeptide

are linked via a linking moiety having Formula I, Formula II or Formula III as defined above.

19. The method according to any one of claims 15 to 18 wherein the method is performed at a pH in the range pH 5.5 to pH 7.5.

20. A method of labelling an oligopeptide, the method comprising the steps:
a) providing a label molecule, the label molecule having a aldehyde or ketone moiety,
b) providing a precursor oligopeptide molecule, the precursor oligopeptide molecule comprising a first oligopeptide fused N-terminally to an intein domain,
c) reacting said precursor oligopeptide molecule with hydrazine to generate an oligopeptide molecule comprising an intermediate oligopeptide, said intermediate oligopeptide having a terminal hydrazide moiety,
d) allowing the aldehyde or ketone moiety of the label molecule to react with the hydrazide moiety of the intermediate oligopeptide molecule to form a labelled oligopeptide product, in which the label molecule and oligopeptide are linked via a hydrazone linking moiety.

21. A labelled oligopeptide produced by the method of any one of claims 15 to 20.

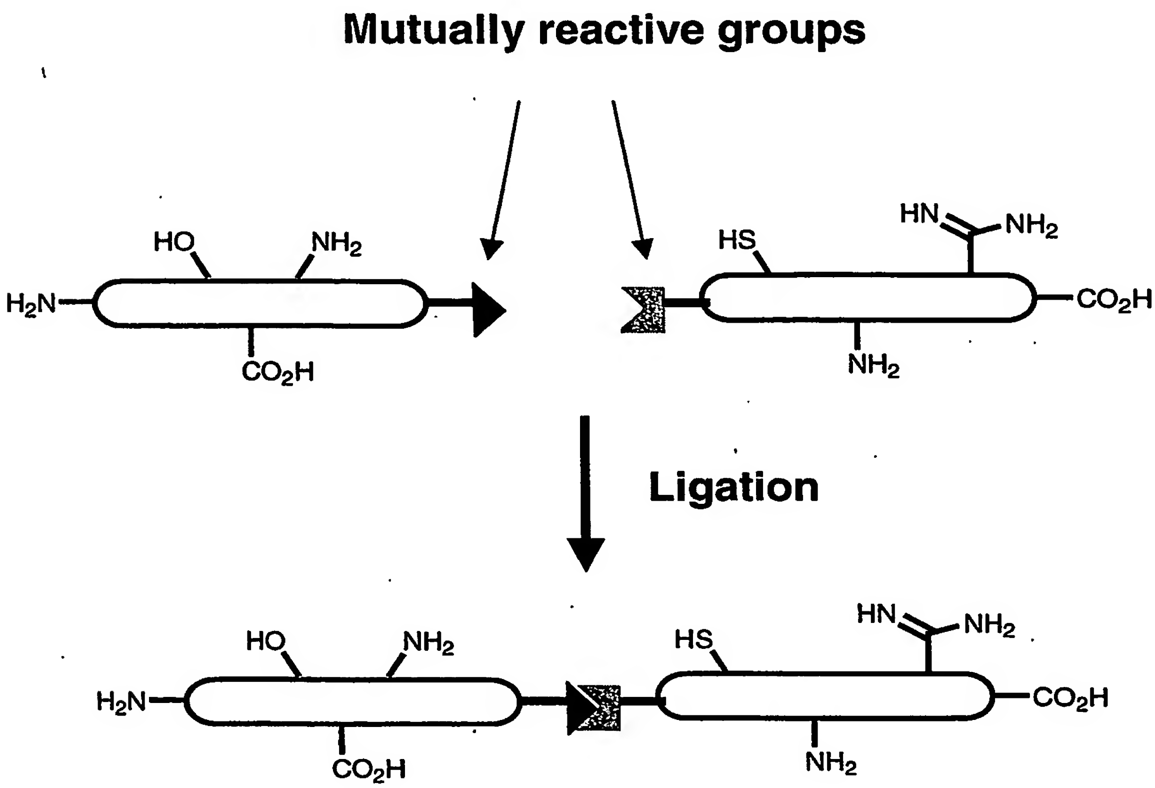


Figure 1 General principle of chemical ligation.

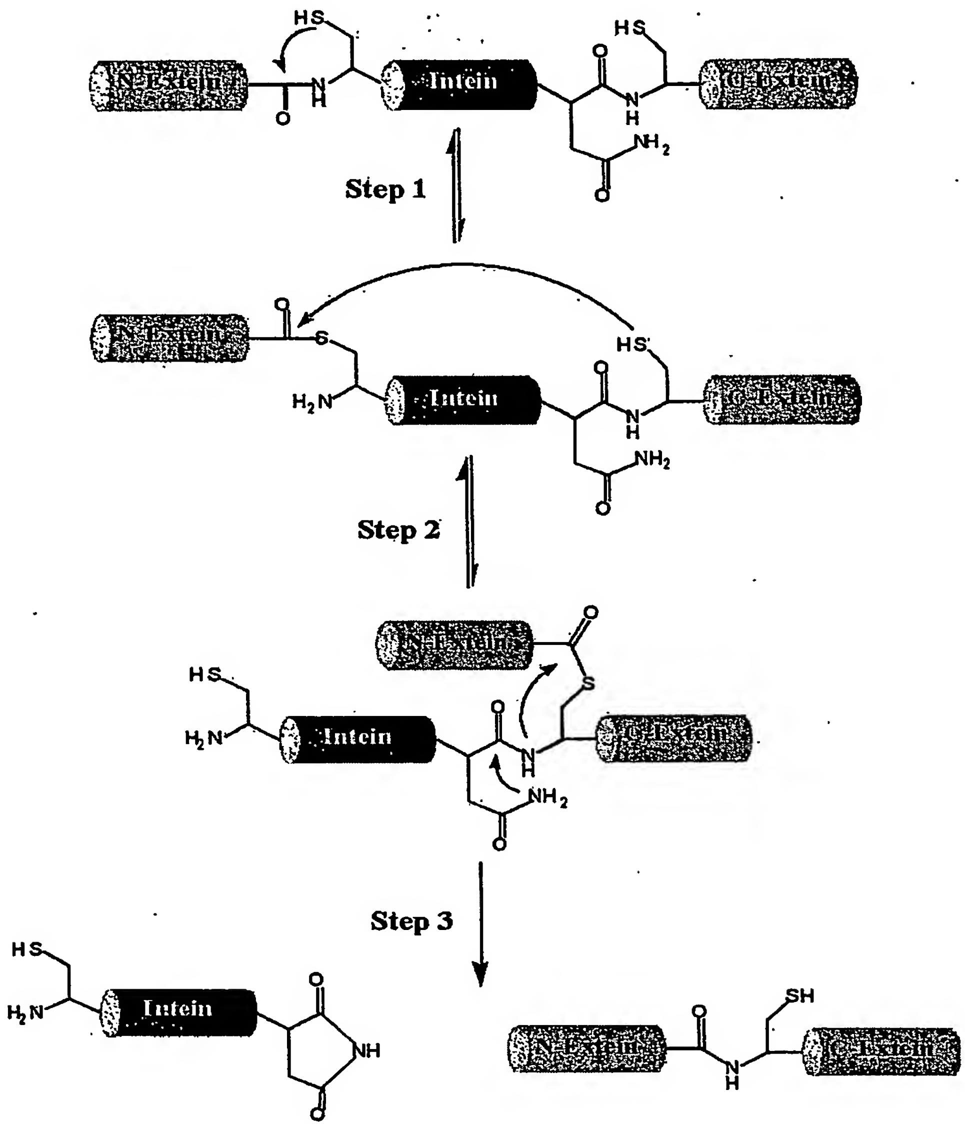


Figure 2 Mechanism of protein splicing

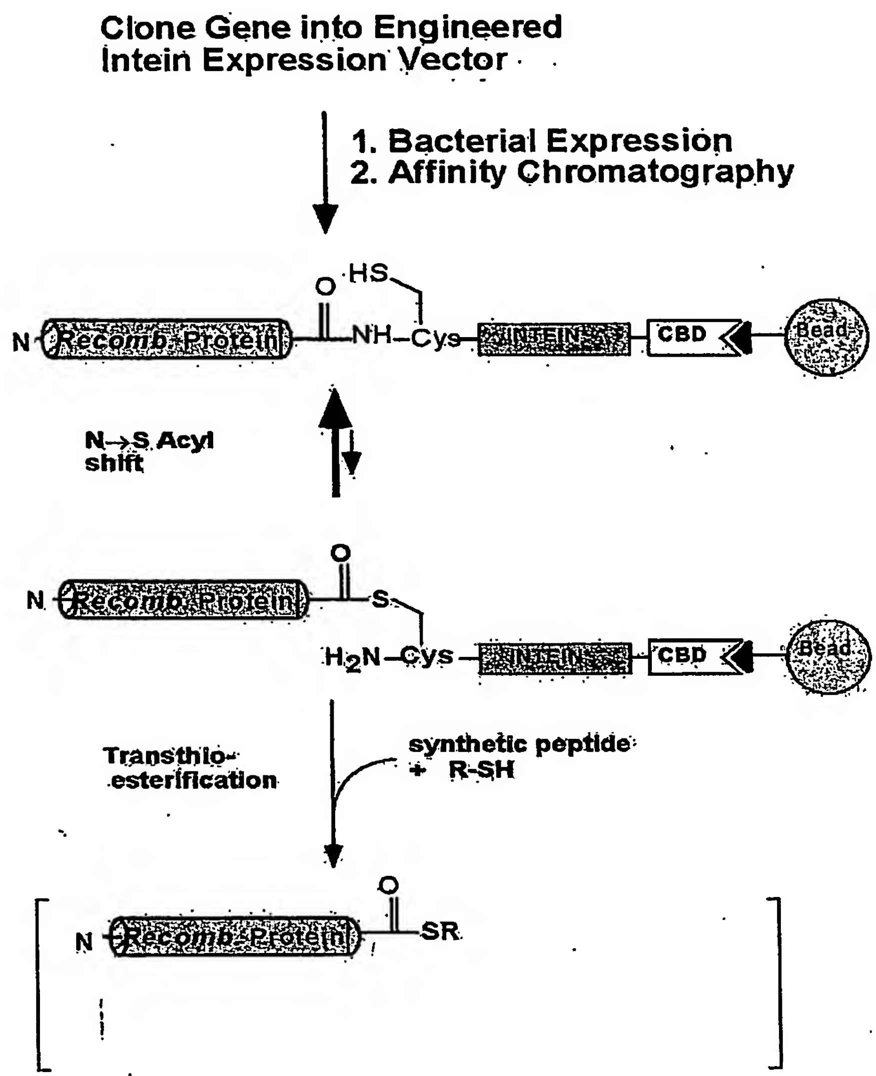


Figure 3 Generation of Recombinant C-terminal Thioester Proteins

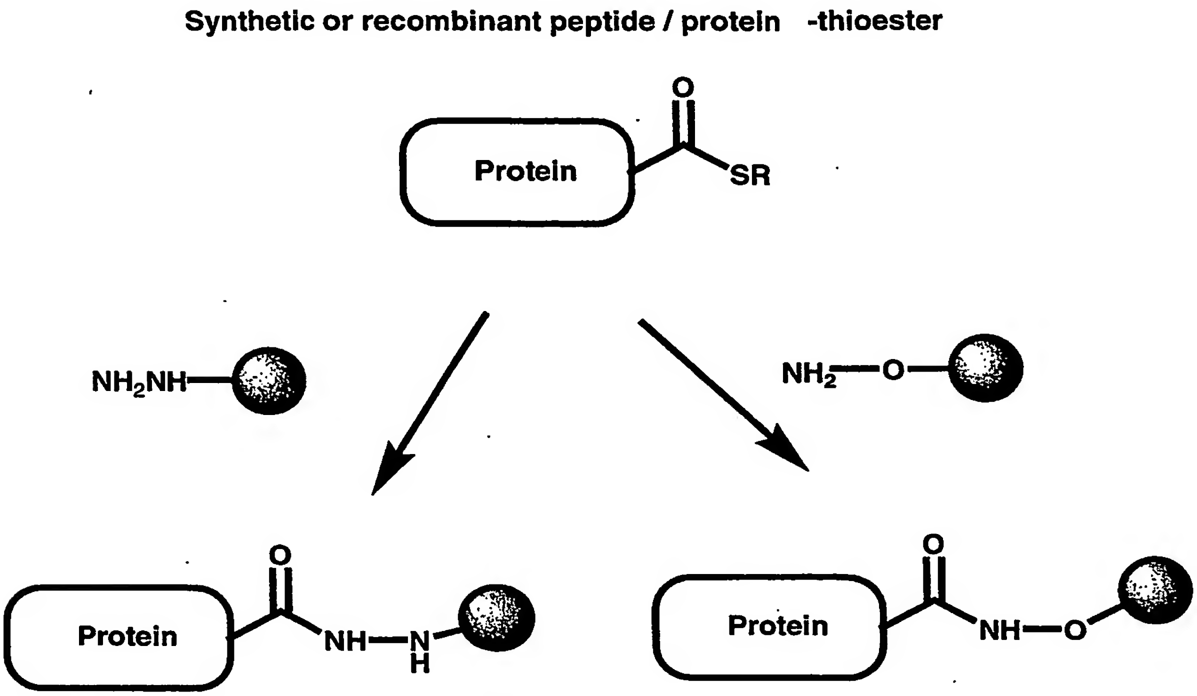


Figure 4 Ligation of protein and peptide thioesters with hydrazine and aminooxy containing entities such as labels, peptides and proteins.

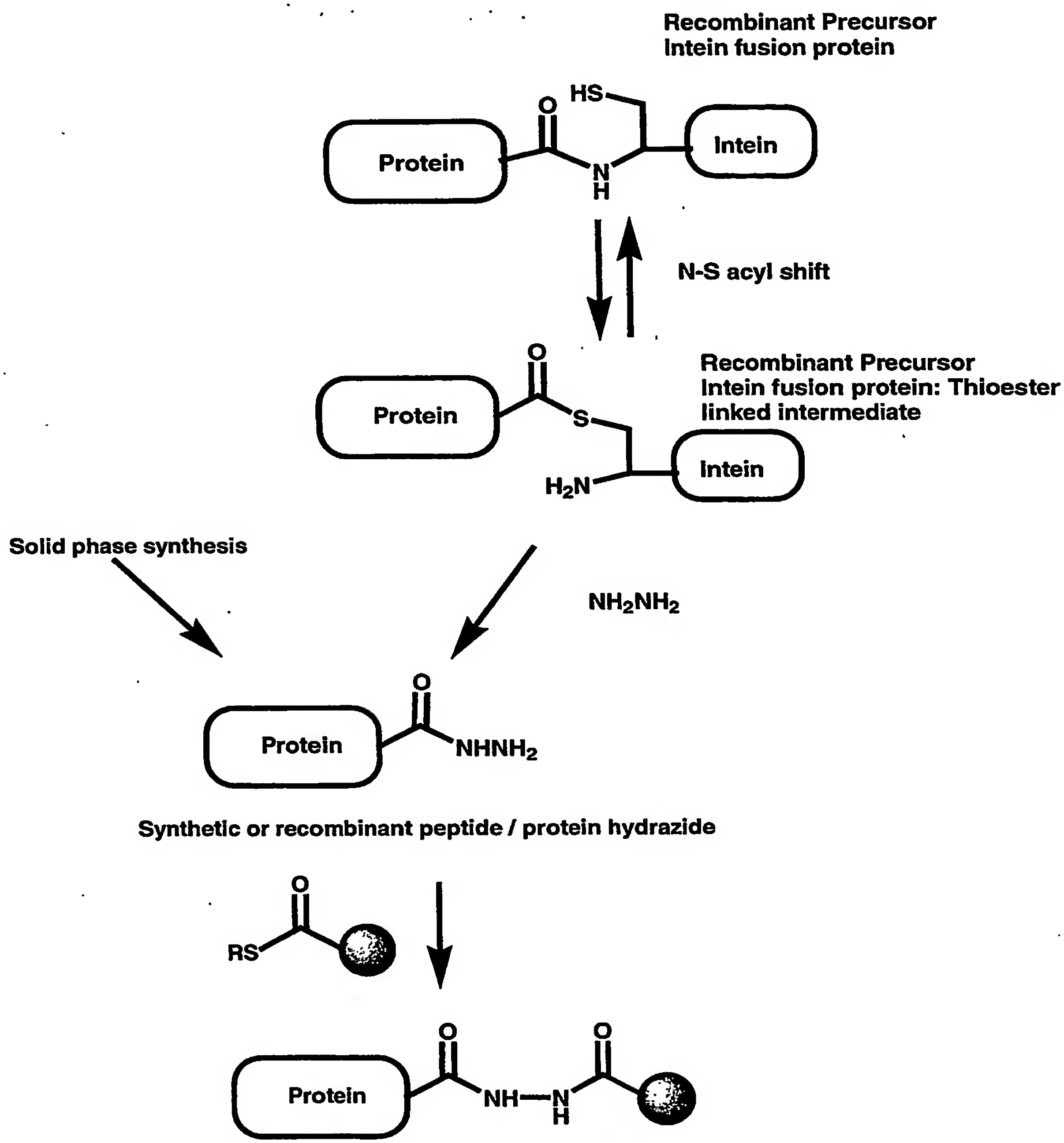


Figure 5 Generation of synthetic and recombinant peptide hydrazides for ligation with thioester containing molecules

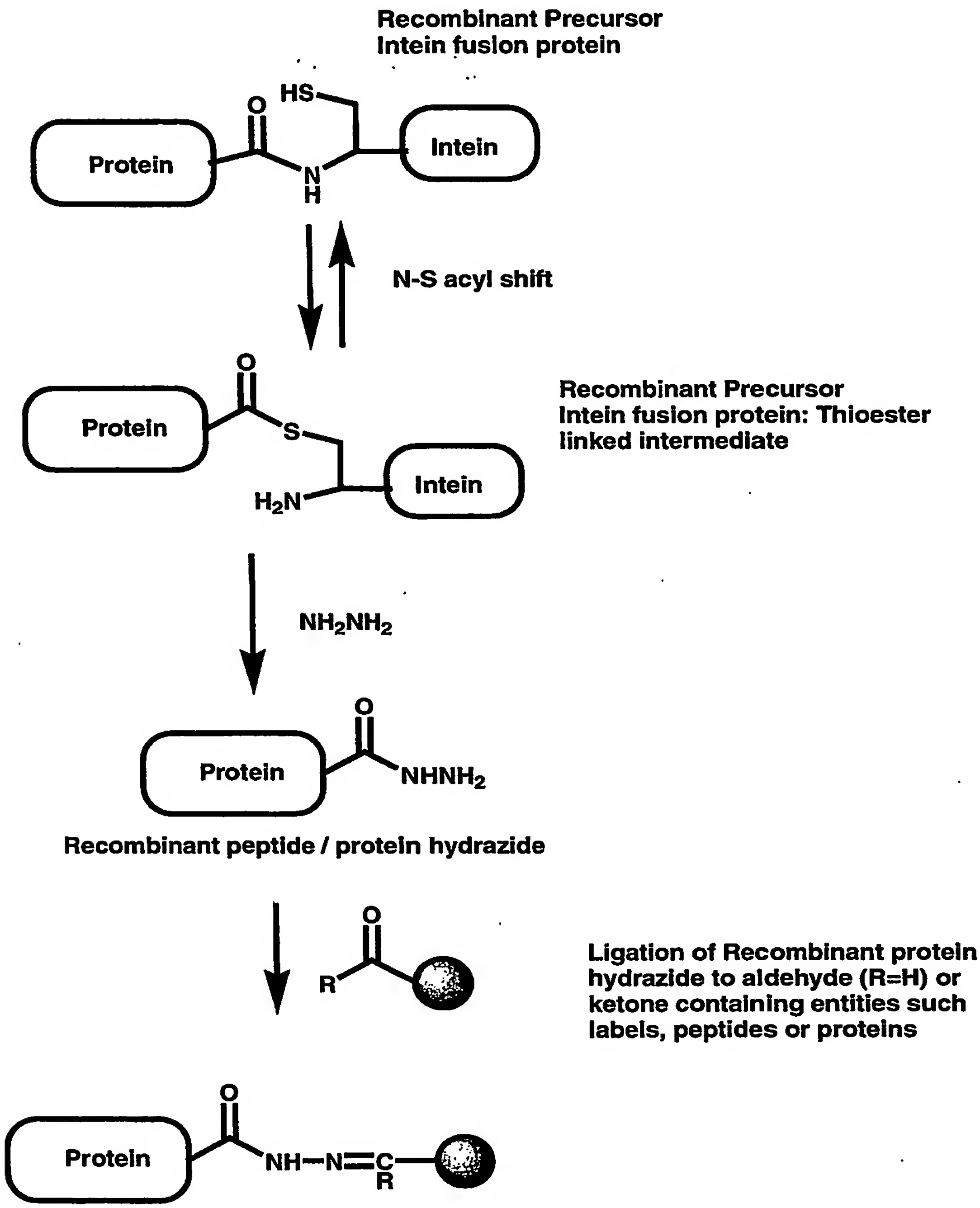


Figure 6 Generation of recombinant peptide hydrazides for ligation with aldehyde and ketone containing molecules

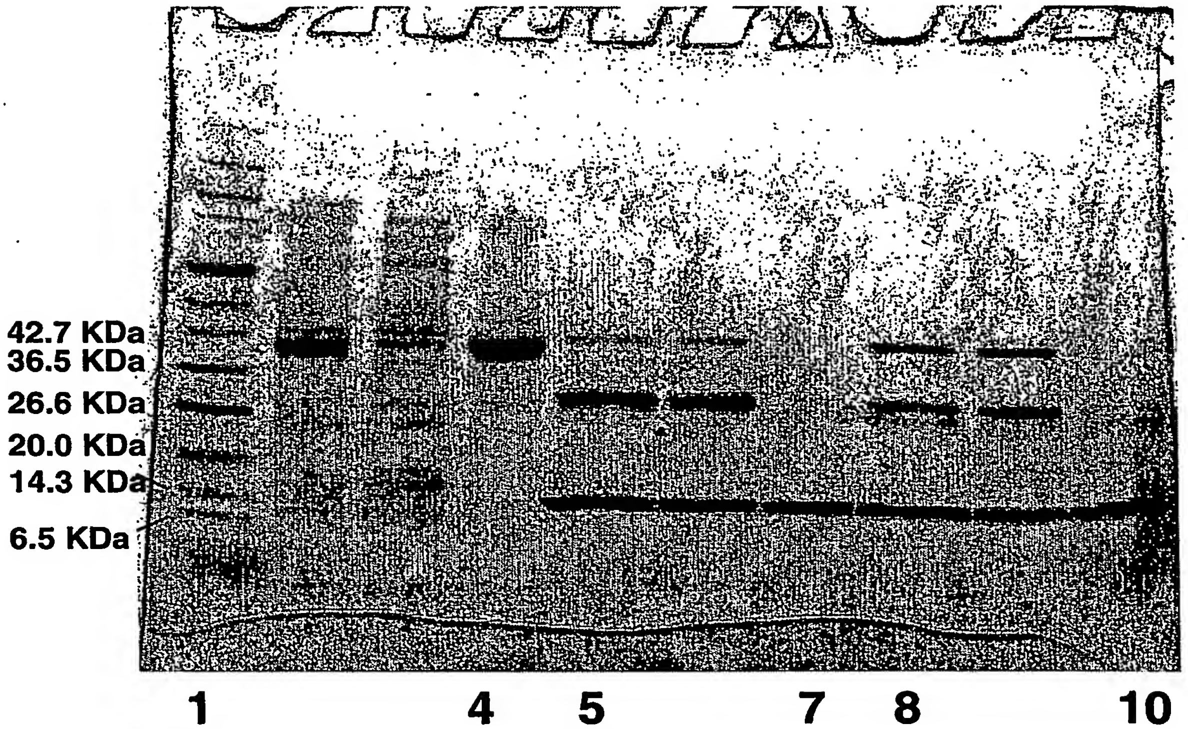


Figure 7. SDS-PAGE analysis of Grb2-SH2 – GyrA – CBD (immobilised on chitin beads) treated with DTT and MESNA. Molecular weight markers (lane 1); purified Grb2-SH2 – GyrA – CBD immobilised on chitin beads (lane 4). Grb2-SH2 – GyrA – CBD treated with 100 mM DTT (lanes 5 and 7) or 120 mM MESNA (lanes 8 and 10). Both the whole reaction slurries (lanes 5 and 8) and the reaction supernatants (lanes 7 and 10) were analysed.

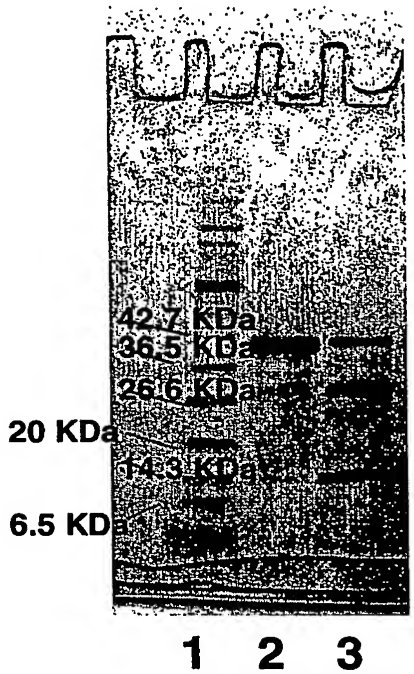


Figure 8. SDS-PAGE analysis of Grb2-SH2 – GyrA – CBD (immobilised on chitin beads) treated with hydrazine. Molecular weight markers (lane 1); Purified Grb2-SH2 – GyrA – CBD immobilised on chitin beads after 20h treatment with phosphate buffer only (lane 2). Grb2-SH2 – GyrA – CBD treated with 200 mM hydrazine in phosphate buffer for 20 h. The whole reaction slurries were analysed.

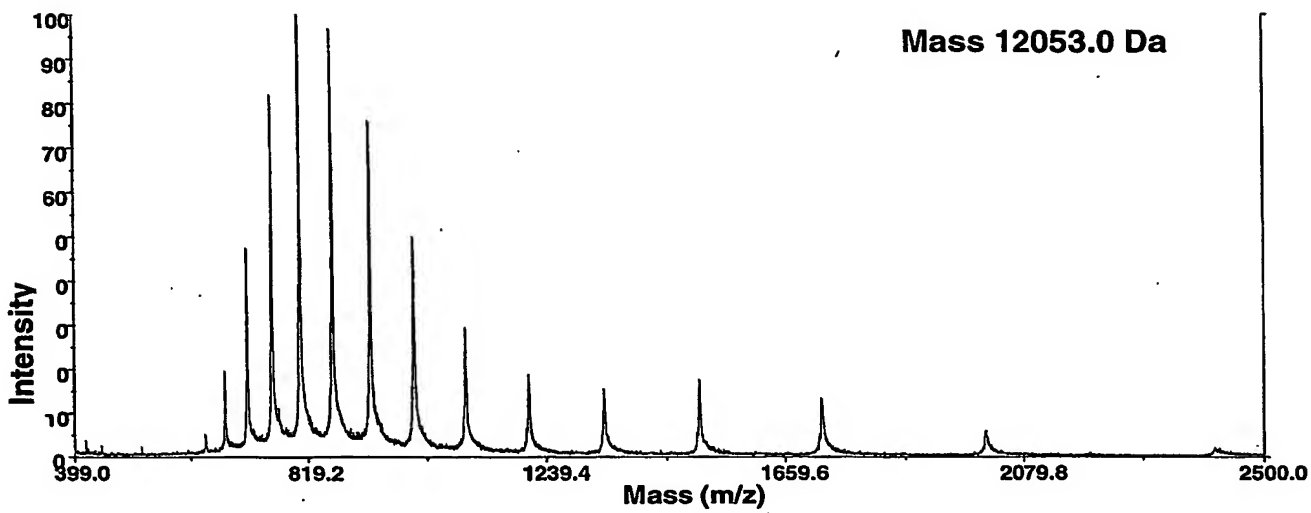


Figure 9. ESMS spectrum of the C-terminal hydrazide derivative of Grb2-SH2

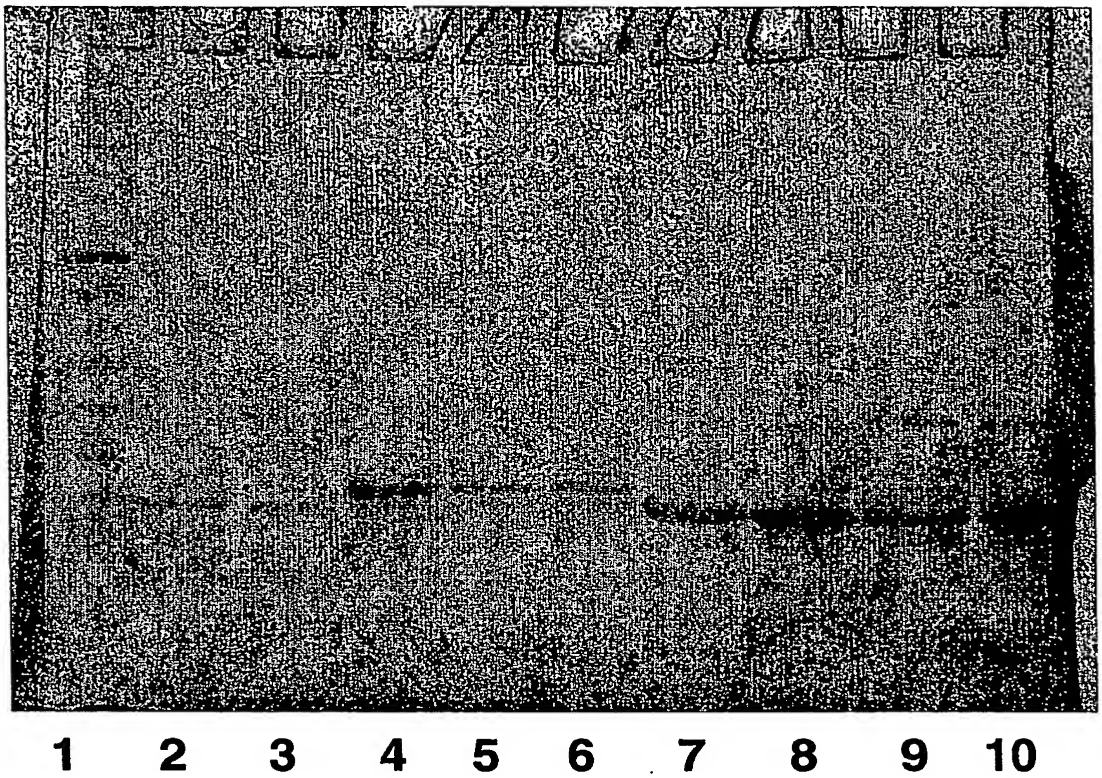


Figure 10. SDS-PAGE analysis of the reaction between synthetic ketone containing peptide CH3COCO-myc with Grb2-SH2 – C-terminal hydrazide and Cytochrome C. Molecular weight markers (lane 1); Grb2-SH2 – C-terminal DTT thioester (lane 2). Reaction between Grb2-SH2 – C-terminal hydrazide and CH3COCO-myc at time points t=0 h (lane 3), t=24 h (lane 4), t= 48h (lane 5) and t= 72 h (lanes 6). Reaction between Cytochrome C and CH3COCO-myc at time points t=0 h (lane 7), t=24 h (lane 8), t= 48h (lane 9) and t= 72 h (lanes 10).